LOCALIZATION ERRORS IN SOLVING STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS IN THE WHOLE SPACE

MÁTÉ GERENCSÉR AND ISTVÁN GYÖNGY

Abstract. Cauchy problems with SPDEs on the whole space are localized to Cauchy problems on a ball of radius R. This localization reduces various kinds of spatial approximation schemes to finite dimensional problems. The error is shown to be exponentially small. As an application, a numerical scheme is presented which combines the localization and the space and time discretization, and thus is fully implementable.

1. Introduction

Parabolic, possibly degenerate, linear stochastic partial differential equations (SPDEs) are considered. In applications such equations are often given on the whole space, which makes the direct implementation of discretization methods problematic. Finite element methods, see e.g., [1], [11], [13], [24], [25] and their references, mostly treat problems on bounded domains and often under strong restrictions on the differential operators, denoted by L and M below. For finite difference schemes convergence results are available on the whole space, see e.g. [26] for the non-degenerate and [2], [7] for the degenerate case, but the schemes themselves are infinite systems of equations. A natural way to overcome this difficulty is to "cut off" the equation outside of a large ball of radius R. A similar approach to obtain error estimates for a truncated terminal condition in deterministic PDEs of optimal stopping problems is used in [23].

The main results of the present paper are Theorems 2.6 and 3.6. Theorem 2.6 compares solutions of two SPDEs whose data agrees on a ball of radius R and establishes an error estimate of order $e^{-\delta R^2}$ in the supremum norm. The proof relies on transforming fully degenerate SPDEs to zero order equations via the method of characteristics, whose analysis goes back to [19] and [18], see also the recent work [21] and the references therein. Once such a result is used to estimate the difference between the original equation and its truncation, one can approximate the truncated equation with known numerical schemes. Our choice is the finite difference method for the spatial and the implicit Euler method for the temporal approximation. The analysis of the former is invoked from [7], while for the latter one requires an error estimate for the time discretization of the finite difference scheme, which is at this point a finite dimensional SDE. Of course this estimate needs to be independent of the spatial mesh size, and this is established in Theorem 3.10. To the authors' best knowledge such an analysis of a full discretization is also a new result for degenerate SPDEs, and in fact even in the deterministic case it improves the results of [2] in

2010 *Mathematics Subject Classification.* Primary 60H15, 60H35, 65M06 .

Key words and phrases. Cauchy problem, degenerate stochastic parabolic PDEs, localization error, finite difference method.

that we are not restricted to finite difference schemes which are monotone. An error estimate for the approximations obtained by localized and fully discretized SPDEs is given in Theorem 3.5, which is a special case of Theorem 3.6, where the accuracy of the approximations is improved by Richardson extrapolation in the spatial discretization.

We mention that an alternative method for localization is to introduce artificial Dirichlet boundary conditions on a large ball, this approach is used for deterministic PDEs in, for example, in [20], [12], and to some extent, in [23]. The order of error in these works is $e^{-\delta R}$, and so the method of the present paper yields an improvement even in the deterministic case. In the present context the method of artificial boundary condition would present additional issues. For instance, if the original SPDE is degenerate, then after introducing the boundary conditions the resulting equation may not even have a solution. Even if we do assume non-degeneracy, in the generality considered here - which is justified and motivated by the filtering equation in signal-observation models - there is no approximation theory of SPDEs on bounded domains with Dirichlet boundary condition, and therefore such a localization method does not help in finding an efficient numerical scheme. One indication of the problem is that regardless of the smoothness of the data and of the boundary, solutions of Dirichlet problems for SPDEs in general do not have continuous second derivatives, see [17].

Let us introduce some notation used throughout the paper. All random elements will be given on a fixed probability space $(\Omega, \mathcal{F}, P)$, equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ of σ-fields $\mathcal{F}_t \subset \mathcal{F}$. We suppose that this probability space carries a sequence of independent Wiener processes $(w^k)_{k=1}^{\infty}$, adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, such that $w^k_t - w^k_s$ is independent of $\mathcal{F}_s$ for each k and any $0 \leq s \leq t$. It is assumed that $\mathcal{F}_0$ contains all P-null subsets of Ω, so that $(\Omega, \mathcal{F}, P)$ is a complete probability space and the σ-fields $\mathcal{F}_t$ are complete. By $\mathcal{P}$ we denote the predictable σ-field of subsets of $\Omega \times [0, \infty)$ generated by $(\mathcal{F}_t)_{t\geq 0}$. We use the shorthand notation $\mathbb{E}^\alpha X = [\mathbb{E}(X)]^\alpha$.

For $p \in [2, \infty)$ and $\vartheta \in \mathbb{R}$ we denote by $L_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$ the space of measurable mappings f from $\mathbb{R}^d$ into a separable Hilbert space $\mathcal{H}$, such that

$$|f|_{L_{p,\vartheta}} = \Big(\int_{\mathbb{R}^d} |(1 + |x|^2)^{\vartheta/2} f(x)|^p_{\mathcal{H}} \, dx \Big)^{1/p} < \infty.$$

We do not include the symbol $\mathcal{H}$ in the notation of the norm in $L_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$. Which $\mathcal{H}$ is involved will be clear from the context. We do the same in other similar situations. In this paper $\mathcal{H}$ will be l_2 or $\mathbb{R}$. The space of functions from $L_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$, whose generalized derivatives up to order m are also in $L_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$, is denoted by $W^m_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$. By definition $W^0_{p,\vartheta}(\mathbb{R}^d, \mathcal{H}) = L_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$. The norm $|f|_{W^m_{p,\vartheta}}$ of f in $W^m_{p,\vartheta}(\mathbb{R}^d, \mathcal{H})$ is defined by

$$|f|^p_{W^m_{p,\vartheta}} = \sum_{|\alpha| \leq m} |D^\alpha f|^p_{L_{p,\vartheta}}, \tag{1.1}$$

where $D^\alpha := D_1^{\alpha_1} ... D_d^{\alpha_d}$ for *multi-indices* $\alpha := (\alpha_1, ..., \alpha_d) \in \{0, 1, ...\}^d$ of *length* $|\alpha| := \alpha_1 + \alpha_2 + ... + \alpha_d$, and $D_i f$ is the generalized derivative of f with respect to x^i for $i = 1, 2..., d$. We also use the notation $D_{ij} = D_i D_j$ and $Df = (D_1 f, ..., D_d f)$. When we talk about "derivatives up to order m" of a function for some nonnegative integer m, then we always include the zeroth-order derivative, i.e. the function

itself. Unless otherwise indicated at some places, the summation convention with respect to repeated integer valued indices is used throughout the paper. The constants in our estimates, usually denoted by N, may change from line to line in the calculations, but their dependencies will always be clear from the statements.

2. Formulation and localization error estimate

Consider the stochastic equation

$$du_t(x) = (L_t u_t(x) + f_t(x))\,dt + \sum_{k=1}^{\infty}(M_t^k u_t(x) + g_t^k(x))\,dw_t^k \tag{2.1}$$

on $(t,x) \in [0,T] \times \mathbb{R}^d =: H_T$, with initial condition

$$u_0(x) = \psi(x), \quad x \in \mathbb{R}^d. \tag{2.2}$$

Here f and $g = (g^k)_{k=1}^{\infty}$ are functions on $\Omega \times H_T$ with values in $\mathbb{R}$ and l_2, respectively, and L and M^k are second order and first order differential operators of the form

$$L_t = a_t^{ij}(x) D_i D_j + b_t^i(x) D_i + c_t(x), \quad M_t^k = \sigma_t^{ik}(x) D_i + \mu_t^k(x), \quad k = 1, 2, ...,$$

where the coefficients a^{ij}, b^i, c, σ^{ik} and μ^k are real-valued functions on $\Omega \times H_T$ for $i, j = 1, 2, ...d$, and integers $k \geq 1$.

For an integer $m \geq 0$, $p \in [2, \infty)$, and $\vartheta \in \mathbb{R}$ the following assumptions ensure the existence and uniqueness of a $W^m_{p,\vartheta}$-valued solution $(u_t(\cdot))_{t\in[0,T]}$.

Assumption 2.1. *For $P \otimes dt \otimes dx$-almost all $(\omega, t, x) \in \Omega \times [0,T] \times \mathbb{R}^d$*

$$\alpha_t^{ij}(x) z^i z^j \geq 0$$

for all $z \in \mathbb{R}^d$, where

$$\alpha^{ij} = 2a^{ij} - \sigma^{ik}\sigma^{jk}.$$

This is a standard assumption in the theory of stochastic PDEs. Below we assume some smoothness on α in x, and on all the coefficients and free terms of problem (2.1)-(2.2). We will also require that the nonnegative symmetric square root $\rho := \sqrt{\alpha}$ possesses bounded second order derivatives in x. Concerning this assumption we remark that it is well-known from [3] that ρ is Lipschitz continuous in x if α is bounded and has bounded second order derivatives, but it is also known that the second order derivatives of ρ may not exist in the classical sense, even if α is smooth with bounded derivatives of arbitrary order.

Assumption 2.2. *(a) The derivatives in $x \in \mathbb{R}^d$ of a^{ij} up to order $\max(m,2)$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i, j \in \{1, 2, ..., d\}$.*

(b) The derivatives in $x \in \mathbb{R}^d$ of b^i and c up to order m are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K for all $i \in \{1, 2, ..., d\}$. The functions $\sigma^i = (\sigma^{ik})_{k=1}^{\infty}$ and $\mu = (\mu^k)_{k=1}^{\infty}$ are l_2-valued and their derivatives in x up to order $m+1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable l_2-valued functions, bounded by K.

(c) The derivatives in $x \in \mathbb{R}^d$ of $\rho = \sqrt{\alpha}$ up to order $m+1$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable functions, bounded by K.

Assumption 2.3. *The initial value, ψ is an $\mathcal{F}_0$-measurable random variable with values in $W^m_{p,\vartheta}$. The free data, f_t and $g_t = (g^k)_{k=1}^{\infty}$ are predictable processes with values in $W^m_{p,\vartheta}$ and $W^{m+1}_{p,\vartheta}(l_2)$, respectively, such that almost surely*

$$\mathcal{K}^p_{m,p,\vartheta}(T) := |\psi|^p_{W^m_{p,\vartheta}} + \int_0^T \left(|f_t|^p_{W^m_{p,\vartheta}} + |g_t|^p_{W^{m+1}_{p,\vartheta}}\right) dt < \infty. \tag{2.3}$$

Definition 2.1. A $W^1_{p,\vartheta}$-valued function u, defined on $[0,T]\times\Omega$, is called a solution of (2.1)-(2.2) on $[0,T]$ if u is predictable on $[0,T]\times\Omega$,

$$\int_0^T |u_t|^p_{W^1_{p,\vartheta}}\,dt < \infty\,(a.s.),$$

and for each $\varphi\in C_0^\infty(\mathbb{R}^d)$ for almost all $\omega\in\Omega$

$$\begin{aligned}(u_t,\varphi) =&(\psi,\varphi)+\int_0^t\{-(a^{ij}_sD_iu_s,D_j\varphi)+(\bar b^i_sD_iu_s+c_su_s+f_s,\varphi)\}\,ds\\ &+\int_0^t(\sigma^{ir}_sD_iu_s+\mu^r_su_s+g^r_s,\varphi)\,dw^r_s\end{aligned}\tag{2.4}$$

for all $t\in[0,T]$, where $\bar b^i=b^i-D_ja^{ij}$, and $(\cdot\,,\cdot)$ denotes the inner product in the Hilbert space of square integrable real-valued functions on $\mathbb{R}^d$.

The following theorem follows from Theorem 2.1 in [6].

Theorem 2.2. *Let Assumptions 2.1, 2.2 (a)-(b), and 2.3 with $m\geq 0$ hold. Then there exists at most one solution on $[0,T]$. If Assumptions 2.1, 2.2(a), and 2.3 hold with $m\geq 1$, then there exists a unique solution $u=(u_t)_{t\in[0,T]}$ on $[0,T]$. Moreover, u is a $W^m_{p,\vartheta}$-valued weakly continuous process, it is a strongly continuous process with values in $W^{m-1}_{p,\vartheta}$, and for every $q>0$ and $n\in\{0,1,...,m\}$*

$$\mathbb{E}\sup_{t\in[0,T]}|u_t|^q_{W^n_{p,\vartheta}}\leq N\mathbb{E}\mathcal{K}^q_{n,p,\vartheta}(T),\tag{2.5}$$

where N is a constant depending only on K, T, d, m, p, ϑ, and q.

Remark 2.3. Theorem 2.1 in [6] covers the $\vartheta=0$ case. We can reduce the case of general ϑ to this by rewriting the equation for u as an equation for $\tilde u(x)=u(x)(1+|x|^2)^{-\vartheta/2}$ (see e.g. [5]). It is easily seen that the coefficients of the resulting equation still satisfy the conditions of the theorem.

Definition 2.4. A $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable random field u on H_T is called a classical solution of (2.1)-(2.2), if along with its derivatives in x up to order 2 it is continuous in $(t,x)\in H_T$, it satisfies (2.1)-(2.2) almost surely for all $(t,x)\in H_T$, and there exists a finite random variable ξ and a constant s such that almost surely

$$|D^\alpha u_t(x)|\leq\xi(1+|x|)^s\quad\text{for all } (t,x)\in H_T \text{ and for } |\alpha|\leq 2.$$

Corollary 2.5. Let assumptions of Theorem 2.2 hold with $m>2+d/p$. Then there exists a unique classical solution u to (2.1)-(2.2).

Let us refer to the problem (2.1)-(2.2) as Eq($\mathfrak{D}$), where $\mathfrak{D}$ stands for the "data"

$$\mathfrak{D}=(\psi,a,b,c,\sigma,\mu,f,g)$$

with $a=(a^{ij})$, $b=(b^i)$, $\sigma=(\sigma^{ki})$, $g=(g^k)$ and $\mu=(\mu^k)$. We are interested in the error when instead of Eq($\mathfrak{D}$) we solve Eq($\bar{\mathfrak{D}}$) with

$$\bar{\mathfrak{D}}=(\bar\psi,\bar a,\bar b,\bar c,\bar\sigma,\bar\mu,\bar f,\bar g).$$

Assumption 2.4. *Almost surely*

$$\mathfrak{D}=\bar{\mathfrak{D}}\quad on\ [0,T]\times\{x\in\mathbb{R}^d:|x|\leq R\}.\tag{2.6}$$

The main example to keep in mind is when each component of $\bar{\mathfrak{D}}$ is a truncation of the corresponding component of $\mathfrak{D}$ (see Section 3 below). Let

$$B_R = \{x \in \mathbb{R}^d : |x| \leq R\}$$

for $R > 0$. Define $\bar{\mathcal{K}}^p_{m,p,\vartheta}(T)$ as $\mathcal{K}^p_{m,p,\vartheta}(T)$ with $\bar{\psi}$, $\bar{f}$ and $\bar{g}$ in place of ψ, f and g, respectively. The main result reads as follows.

Theorem 2.6. *Let $\nu \in (0,1)$ and let Assumptions 2.1, 2.2 (b)-(c) and 2.3 hold with $m > 2 + d/p$ and $\vartheta \in \mathbb{R}$ for $\mathfrak{D}$ and $\bar{\mathfrak{D}}$. Let also Assumption 2.4 hold. Then Eq($\mathfrak{D}$) and Eq($\bar{\mathfrak{D}}$) have a unique classical solution u and $\bar{u}$, respectively, and for $q > 0$, $r > 1$*

$$\mathbb{E} \sup_{t\in[0,T]} \sup_{x\in B_{\nu R}} |u_t(x) - \bar{u}_t(x)|^q \leq N e^{-\delta R^2} \mathbb{E}^{1/r}(\mathcal{K}^{qr}_{m,p,\vartheta}(T) + \bar{\mathcal{K}}^{qr}_{m,p,\vartheta}(T)), \tag{2.7}$$

where N and δ are positive constants, depending on K, d, T, q, r, ϑ, p, and ν.

First we collect some auxiliary results. The following lemma is a version of Kolmogorov's continuity criterion, see Theorem 3.4 of [4].

Lemma 2.7. *Let $x(\theta)$ be a stochastic process parametrized by and continuous in $\theta \in D \subset \mathbb{R}^p$, where D is a direct product of lower dimensional closed balls. Then for all $0 < \alpha < 1$, $q \geq 1$, and $s > p/\alpha$,*

$$\mathbb{E}\sup_\theta |x(\theta)|^q \leq N(1+|D|)\left[\sup_\theta (\mathbb{E}|x(\theta)|^{qs})^{1/s} + \sup_{\theta\neq\theta'}\left(\frac{\mathbb{E}|x(\theta)-x(\theta')|^{qs}}{|\theta-\theta'|^{qs\alpha}}\right)^{1/s}\right]$$

where $N = N(q,s,\alpha,p)$, and $|D|$ is the volume of D.

Lemma 2.8. *Let $(\alpha_t)_{t\in[0,T]}$ and $(\beta_t)_{t\in[0,T]}$ be $\mathcal{F}_t$-adapted processes with values in $\mathbb{R}^d$ and $l_2(\mathbb{R}^d)$, respectively, in magnitude bounded by a constant K. Then for the process*

$$X_t = \int_0^t \alpha_s\,ds + \int_0^t \beta^k_s\,dw^k_s, \quad t \in [0,T] \tag{2.8}$$

there exist constants $\varepsilon = \varepsilon(K,T) > 0$ and a $N = N(K,T)$ such that

$$\mathbb{E}\sup_{t\leq T} e^{\varepsilon|X_t|^2} \leq N.$$

Proof. A somewhat more general lemma is proved in [23]. For convenience of the reader we give the proof here. By Itô's formula

$$Y_t := e^{|X_t|^2 e^{-\mu t}} = 1 + \int_0^t e^{|X_s|^2 e^{-\mu s} - \mu s}\{|\beta_s|^2 + 2\alpha_s X_s$$
$$+2|\beta_s X_s|^2 - \mu|X_s|^2\}\,ds + m_t$$

for any $\mu \in \mathbb{R}$, where $(m_t)_{t\in[0,T]}$ is a local martingale starting from 0. By simple inequalities

$$2\alpha X + 2|\beta X|^2 \leq |\alpha|^2 + |X|^2 + 2|\beta|^2|X|^2 \leq K^2 + (2K^2+1)|X|^2.$$

Hence for $\mu = (2K^2+1)$ and for a stopping time $\tau \leq T$ we have

$$\mathbb{E}Y_{t\wedge\tau_n} \leq 1 + 2K^2\int_0^t \mathbb{E}Y_{s\wedge\tau_n}\,ds,$$

for $\tau_n = \tau \wedge \rho_n$, where $(\rho_n)_{n=1}^{\infty}$ is a localizing sequence of stopping times for m. Hence, by Gronwall's lemma,

$$\mathbb{E}Y_{t\wedge\tau_n} \leq e^{2K^2T}.$$

where N is independent of n. Letting here $n \to \infty$, by Fatou's lemma we get

$$\mathbb{E}e^{|X_\tau|^2 e^{-\mu T}} \leq \mathbb{E}e^{|X_\tau|^2 e^{-\mu\tau}} \leq e^{K^2T}$$

for stopping times $\tau \leq T$. Hence applying Lemma 3.2 from [8] for $r \in (0,1)$ we obtain

$$\mathbb{E}\sup_{t\leq T} e^{r|X_\tau|^2 e^{-\mu T}} \leq \tfrac{2-r}{1-r} e^{rK^2T}.$$

□

To formulate our next lemma we consider the stochastic differential equation

$$dX_s = \alpha_s(X_s)\,ds + \beta_s^k(X_s)\,dw_s^k, \tag{2.9}$$

where α and $\beta = (\beta^k)$ are $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$-measurable function on $\Omega\times[0,T]\times\mathbb{R}^d$, with values in $\mathbb{R}^d$ and $l_2(\mathbb{R}^d)$ such that they are bounded in magnitude by K and satisfy the Lipschitz condition in $x \in \mathbb{R}^d$ with a Lipschitz constant M, uniformly in the other arguments. Then equation (2.9) with initial condition $X_t = x$ has a unique solution $X^{t,x} = (X_s^{t,x})_{s\in[t,T]}$ for any $t \in [0,T]$ and $x \in \mathbb{R}^d$.

Remark 2.9. It is well known from [19] that the solution of (2.8) can be chosen to be continuous in t, x, s. In the following, by $X_s^{t,x}$ we always understand such a continuous modification.

Lemma 2.10. *Set $\hat{X}^{t,x} = X^{t,x} - x$. There exists a constant $\delta = \delta(d,K,M,T) > 0$ such that for any R,*

$$\mathbb{E}\sup_{0\leq t\leq s\leq T}\sup_{|x|\leq R} e^{|\hat{X}_s^{t,x}|^2\delta} \leq N(1+R^{d+1/2}), \tag{2.10}$$

and for any R and r

$$P(\sup_{0\leq t\leq s\leq T}\sup_{|x|\leq R}|\hat{X}_s^{t,x}| > r) \leq Ne^{-\delta r^2}(1+R^{d+1/2}), \tag{2.11}$$

where $N = N(d,K,M,T)$.

Proof. It is easy to see that (2.10) implies (2.11), so we need only prove the former. For a fixed δ, to be chosen later, let us use the notations $f(y) = e^{|y|^2\delta}$ and $\gamma = 2(d+2)+1$. By Lemma 2.7, we have

$$\mathbb{E}\sup_{0\leq t\leq s\leq T}\sup_{|x|\leq R} f(\hat{X}_s^{t,x}) \leq N(1+R^d)\sup_{0\leq t\leq s\leq T}\sup_{|x|\leq R}(\mathbb{E}f^\gamma(\hat{X}_s^{t,x}))^{1/\gamma}$$

$$+N(1+R^d)\sup_{\substack{0\leq t\leq s\leq T\\ 0\leq t'\leq s'\leq T}}\sup_{\substack{|x|\leq R\\ |x'|\leq R}}\left(\frac{\mathbb{E}|f(\hat{X}_s^{t,x}) - f(\hat{X}_{s'}^{t',x'})|^\gamma}{(|t-t'|^2+|s-s'|^2+|x-x'|^2)^{\gamma/4}}\right)^{1/\gamma}. \tag{2.12}$$

The first term above, by Lemma 2.8, provided $\delta \leq \varepsilon/\gamma$, can be estimated by NR^d. As for the second one,

$$f(\hat{X}_s^{t,x}) - f(\hat{X}_{s'}^{t',x'}) = \int_0^1 \partial f(\vartheta\hat{X}_s^{t,x} + (1-\vartheta)\hat{X}_s^{t',x'})(\hat{X}_s^{t,x} - \hat{X}_{s'}^{t',x'})\,d\vartheta.$$

Notice that $|\nabla f(y)| \leq N(\delta) f^2(y)$, therefore by Jensen's inequality and Lemma 2.8 again, provided $\delta \leq \varepsilon/(8\gamma)$, we obtain

$$\mathbb{E}|f(\hat{X}_s^{t,x}) - f(\hat{X}_{s'}^{t',x'})|^\gamma \leq N \mathbb{E}^{1/2} |\hat{X}_s^{t,x} - \hat{X}_{s'}^{t',x'}|^{2\gamma}.$$

Now the the right-hand side can be estimated by standard moment bounds for SDEs, see e.g. Corollary 2.5.5 in [15], from which we obtain

$$\left(\frac{\mathbb{E}|f(\hat{X}_s^{t,x}) - f(\hat{X}_{s'}^{t',x'})|^{2\gamma}}{(|t-t'|^2 + |s-s'|^2 + |x-x'|^2)^{\gamma/2}} \right)^{1/(2\gamma)} \leq N(1 + R^{1/2}).$$

□

Proof of Theorem 2.6.

Throughout the proof we will use the constant $\lambda = \lambda(d, q)$, which stands for a power of R, and, like N and δ, may change from line to line. Clearly it suffices to prove Theorem 2.6 with $e^{-\delta R^2} R^\lambda$ in place of $e^{-\delta R^2}$ in the right-hand side of inequality (2.7). We also assume first that $q > 1$ and $\vartheta = 0$.

The main idea of the proof is based on stochastic representation of solutions to linear stochastic PDEs of parabolic type, see [18], [19], and [21]. This representation can be viewed as the generalization of the well-known Feynman-Kac formula and is derived as follows. First, we consider an equation which differs from the original only by an additional stochastic term driven by an independent Wiener process. The new equation is fully degenerate and taking conditional expectation with respect to the original filtration of its solution gives back u. On the other hand, the method of characteristics allows us to transform the fully degenerate equation to a much simpler one. This provides a formula for the representation of u, and, more importantly for our purposes, allows us to compare u and $\bar{u}$ on the level of characteristics.

Recall that $\rho = (\rho_t^{ir}(x))_{i,r=1}^d$ is the symmetric nonnegative square root of $\alpha = (2a^{ij} - \sigma^{ik}\sigma^{jk})_{i,j=1}^d$ and $\bar{\rho}$ is the symmetric nonnegative square root of $\bar{\alpha} = (2\bar{a}^{ij} - \bar{\sigma}^{ik}\bar{\sigma}^{jk})_{i,j=1}^d$. Then due to Assumption 2.4, $\rho = \bar{\rho}$ almost surely for all $t \in [0, T]$ and for $|x| \leq R$. Let $(\hat{w}_t^r)_{t \geq 0, r=1\dots d}$ be a d-dimensional Wiener process, also independent of the σ-algebra $\mathcal{F}_\infty$ generated by $\mathcal{F}_t$ for $t \geq 0$. Consider the problem

$$\begin{aligned} dv_t(x) =& (L_t v_t(x) + f_t(x))\, dt + (M_t^k v_t(x) + g_t^k(x))\, dw_t^k \\ &+ \mathcal{N}_t^r v_t(x)\, d\hat{w}_t^r \end{aligned} \tag{2.13}$$

$$v_0(x) = \psi(x), \tag{2.14}$$

where $\mathcal{N}^r = \rho^{ri} D_i$. Then by Corollary 2.5, (2.13)-(2.14) has a unique classical solution v, and for each $t \in [0, T]$ and $x \in \mathbb{R}^d$ almost surely

$$u_t(x) = \mathbb{E}(v_t(x)|\mathcal{F}_t). \tag{2.15}$$

Together with (2.13) let us consider the stochastic differential equation

$$dY_t = \beta_t(Y_t)\, dt - \sigma_t^k(Y_t)\, dw_t^k - \rho_t^r(Y_t)\, d\hat{w}_t^r, \quad 0 \leq t \leq T, \quad Y_0 = y, \tag{2.16}$$

where

$$\beta_t(y) = -b_t(y) + \sigma_t^{ik}(y) D_i \sigma_t^k(y) + \rho_t^{ri}(y) D_i \rho_t^r(y) + \sigma_t^k(y) \mu_t^k(y), \quad t \in [0, T],\, y \in \mathbb{R}^d,$$

and σ^k, ρ^r stand for the column vectors $(\sigma^{1k},\dots,\sigma^{dk})$, $(\rho^{1r},\dots,\rho^{dr})$, respectively. By the Itô-Wentzell formula from [16], for

$$U_t(y) := v_t(Y_t(y))$$

we have (to ease the notation we omit the parameter y in $Y_t(y)$)

$$\begin{aligned} dv_t(Y_t) &= (L_t v_t(Y_t) + f_t(Y_t))\,dt + (M_t^k v_t(Y_t) + g_t^k(Y_t))\,dw_t^k + \mathcal{N}_t^r v_t(Y_t)\,d\hat{w}_t^r \\ &\quad + (\beta_t^i D_i v_t(Y_t) + a_t^{ij} D_{ij} v_t(Y_t))\,dt - \sigma_t^{ik} D_i v_t(Y_t)\,dw_t^k - \mathcal{N}_t^r v(Y_t)\,d\hat{w}_t^r \\ &\quad - \sigma_t^{ik} D_i(M_t^k v_t(Y_t) + g^k(Y_t))\,dt - \mathcal{N}_t^r \mathcal{N}_t^r v_t(Y_t)\,dt. \end{aligned} \tag{2.17}$$

Due to cancellations on the right-hand side of (2.17) we obtain

$$\begin{aligned} dU_t(y) =& \{\gamma_t(Y_t(y))U_t(y) + \phi_t(Y_t(y))\}\,dt \\ &+ \{\mu_t^k(Y_t(y))U_t(y) + g_t^k(Y_t(y))\}\,dw_t^k, \quad U_0(y) = \psi(y), \end{aligned}$$

where

$$\gamma_t(x) := c_t(x) - \sigma_t^{ki}(x) D_i \mu_t^k(x), \quad \phi_t(x) = f_t(x) - \sigma_t^{ki}(x) D_i g_t^k.$$

Notice that in the special case when $f = 0$, $g = 0$, $c = 0$, $\mu = 0$ and $\psi(x) = x^i$ for $i \in \{1,...,d\}$, we get $\tilde{v}_t^i(Y_t(y)) = y^i$ for $i = 1,...,d$, where $\tilde{v}^i$ is the solution of (2.13)-(2.14) with $f = c = 0$, $g = \mu = 0$, $\sigma = 0$ and $\psi(x) = x^i$. Hence for each $t \in [0,T]$ the mapping $y \to Y_t(y) \in \mathbb{R}^d$ has an inverse, Y_t^{-1}, for almost every ω, and the mapping $x \to \tilde{v}_t(x) = (\tilde{v}_t^i(x))_{i=1}^d$, defined by the continuous random field $(\tilde{v}_t^i)_{(t,x)\in H_T}$ gives a continuous modification of Y_t^{-1}. Also, we can write $v_t(x) = U_t(Y_t^{-1})$.

As we shall see, due to the data being the same on a large ball, the characteristics Y and $\bar{Y}$ agree on an event of large probability. This fact and the above representation will yield the estimate (2.7). Set $\bar{U}_t(y) = \bar{v}_t(\bar{Y}_t(y))$, where $\bar{v}_t(x)$ and $\bar{Y}_t(y)$ are defined as $v_t(x)$ and $Y_t(y)$ in (2.13)-(2.14) and (2.16), respectively, with $\bar{\mathfrak{D}}$ and $\bar{\rho}$ in place of $\mathfrak{D}$ and ρ.

Introduce the notations $\mathbf{B}_R = [0,T] \times B_R$ and $\mathbf{A}_R = \mathbf{B}_R \cap \mathbb{Q}^{d+1}$. Since u and $\bar{u}$ are continuous in both variables,

$$\sup_{(t,x)\in\mathbf{B}_{\nu R}} |u_t(x) - \bar{u}_t(x)| = \sup_{(t,x)\in\mathbf{A}_{\nu R}} |u_t(x) - \bar{u}_t(x)| \tag{2.18}$$

Let $\nu' = (1+\nu)/2$ and define the event

$$H := \left[\sup_{(t,x)\in\mathbf{B}_{\nu R}} |Y_t^{-1}(x)| > \nu' R\right] \cup \left[\sup_{(t,x)\in\mathbf{B}_{\nu' R}} |Y_t(x)| > R\right].$$

Then

$$H^c = \left[Y_t^{-1}(x) \in B_{\nu' R}, \forall (t,x) \in \mathbf{B}_{\nu R}\right] \cap \left[Y_t(x) \in B_R, \forall (t,x) \in \mathbf{B}_{\nu' R}\right],$$

and thus on H^c

$$\begin{aligned} Y_t(x) &= \bar{Y}_t(x) \quad \text{for } (t,x) \in \mathbf{B}_{\nu' R}, \\ Y_t^{-1}(x) &= \bar{Y}_t^{-1}(x) \quad \text{for } (t,x) \in \mathbf{B}_{\nu R}\ , \end{aligned}$$

and consequently,

$$v_t(x) = \bar{v}_t(x) \quad \text{for } (t,x) \in \mathbf{B}_{\nu R}.$$

Therefore, by (2.15) and (2.18), and by Doob's, Hölder's, and the conditional Jensen inequalities,

$$\mathbb{E} \sup_{(t,x)\in\mathbf{B}_{\nu R}} |u_t(x) - \bar{u}_t(x)|^q \le \mathbb{E} \sup_{t\in[0,T]\cap\mathbb{Q}} |\mathbb{E}(\mathbf{1}_H \sup_{(\tau,x)\in\mathbf{A}_{\nu R}} |v_\tau(x) - \bar{v}_\tau(x)| | \mathcal{F}_t)|^q$$

$$\le \frac{q}{q-1}(P(H))^{1/r'}\mathbb{E}^{1/r}\Big(\sup_{(\tau,x)\in H_T}|v_\tau(x)-\bar v_\tau(x)|^{qr'}\Big) \tag{2.19}$$

$$\le \frac{2^{q-1}q}{q-1}(P(H))^{1/r'}V_T \tag{2.20}$$

with

$$V_T := \mathbb{E}^{1/r}\sup_{(\tau,x)\in H_T}|v_\tau(x)|^{qr} + \mathbb{E}^{1/r}\sup_{(\tau,x)\in H_T}|\bar v_\tau(x)|^{qr},$$

for $r>1$, $r'=r/(r-1)$, provided $q>1$. By Theorem 2.2

$$V_T \le N\mathbb{E}^{1/r}(\mathcal{K}^{qr}_{m,p,0}(T)+\bar{\mathcal{K}}^{qr}_{m,p,0}(T)). \tag{2.21}$$

We can estimate $P(H)$ as follows. Clearly,

$$P(H) \le P\Big(\sup_{(t,x)\in\mathbf{B}_{\nu R}}|Y_t^{-1}(x)|>\nu' R\Big)+P\Big(\sup_{(t,x)\in\mathbf{B}_{\nu' R}}|Y_t(x)|>R\Big) =: J_1+J_2.$$

For $\hat Y_t(x)=Y_t(x)-x$ by (2.11) we have

$$J_2 \le P\Big(\sup_{(t,x)\in\mathbf{B}_{\nu' R}}|\hat Y_t(x)|>(1-\nu')R\Big) \le NR^{d+1/2}e^{-\delta(1-\nu)^2R^2}.$$

Also, we have

$$J_1 \le \sum_{l=0}^\infty P(\exists(t,x)\in[0,T]\times(B_{2^{l+1}\nu' R}\setminus B_{2^l\nu' R}) : |Y_t(x)|\le \nu R)$$

$$\le \sum_{l=0}^\infty P\Big(\sup_{(t,x)\in\mathbf{B}_{2^{l+1}\nu' R}}|\hat Y_t(x)|\ge(2^l\nu'-\nu)R\Big).$$

Using (2.11) again gives

$$J_1 \le N\sum_{l=0}^\infty e^{-\delta(2^l\nu'-\nu)^2R^2}(2^{l+1}\nu' R)^{d+1} \le Ne^{-\delta R^2}$$

We can conclude that

$$P(H) \le Ne^{-\delta R^2}, \tag{2.22}$$

where N and δ are positive constants, depending only on d, K and T.

Combining this with (2.20) and (2.21) we can finish the proof of the theorem under the additional conditions.

For general ϑ one applies the same arguments as in Remark 2.3. Finally (2.7) for the case $q\in(0,1]$ follows easily from standard arguments using Lemma 3.2 from [8]. □

3. An application - finite differences

In this section we apply Theorem 2.6 to present a numerical scheme approximating the initial value problem (2.1)-(2.2). We make use of the results of [7] on the rate and acceleration of finite difference approximations in the spatial variable, which, together with a time discretization and a truncation - whose error can be estimated using Theorem 2.6 - yields a fully implementable scheme. We shall carry out the steps of approximation in the following order: spatial discretization by finite differences, localization of the finite difference scheme, and discretization in time via implicit Euler's method. This of course requires an analysis of the Euler scheme, to present an error estimate for it, which does not depend on the spatial

mesh size and the localization. Furthermore, in our full discretization scheme we shall incorporate Richardson's extrapolation, which will allow us to improve the accuracy of the scheme in the spatial mesh size h.

First we introduce the finite difference approximation in the spatial variable for (2.1)-(2.2). To this end, let $\Lambda_1 \subset \mathbb{R}^d$ be a finite set, containing the zero vector, satisfying the following natural condition: $\Lambda_0 := \Lambda_1 \setminus \{0\}$ is not empty, and if a subset $\Lambda' \subset \Lambda_0$ is linearly dependent, then it is linearly dependent over the rationals. Let $h > 0$, and define the grid

$$\mathbb{G}_h = \{h \sum_{i=1}^{n} \lambda_i : \lambda_i \in \Lambda_1 \cup -\Lambda_1, n = 1, 2, \ldots\}.$$

Due to the assumption on Λ_1, $\mathbb{G}_h$ has only finitely many points in every ball around the origin in $\mathbb{R}^d$. Define for $\lambda \in \Lambda_0 \cup -\Lambda_0$, the finite difference operators

$$\delta^h_\lambda \varphi(x) = \frac{1}{2h}(\varphi(x + h\lambda) - 2\varphi(x) + \varphi(x - h\lambda)),$$

and let δ^h_0 stand for the identity operator. To approximate the Cauchy problem (2.1) -(2.2), for $h > 0$ we consider the equation

$$du_t(x) = (L^h_t u_t(x) + f_t(x))\, dt + \sum_{k=1}^{\infty}(M^{h,k}_t u_t(x) + g^k_t(x))dw^k_t \tag{3.1}$$

on $[0, T] \times \mathbb{G}_h$, with initial condition

$$u_0(x) = \psi(x), \tag{3.2}$$

where L^h and $M^{h,k}$ are difference operators of the form

$$L^h_t(x) = \sum_{\lambda,\kappa\in\Lambda_1} \mathfrak{a}^{\lambda\kappa}_t(x)\delta^h_\lambda\delta^h_\kappa, \quad M^{h,k}_t(x) = \sum_{\lambda\in\Lambda_1} \mathfrak{b}^{\lambda,k}_t(x)\delta^h_\lambda, \quad k = 1, 2, \ldots,$$

with some real-valued $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$-measurable $\mathfrak{a}^{\lambda\kappa}_t$ and $\mathfrak{b}^{\lambda,k}$ on $\Omega \times [0, T]$, such that

$$|\mathfrak{a}^{\lambda\kappa}_t(x)| \le K \quad \text{and} \quad \sum_k |\mathfrak{b}^{\lambda,k}_t(x)|^2 \le K^2 \tag{3.3}$$

for all $\lambda, \kappa \in \Lambda_1$, $t \in [0, T]$, $x \in \mathbb{R}^d$ and $\omega \in \Omega$, where K is a constant.

Remark 3.1. Here ψ, f and g are the same as in (2.1)-(2.2) and we will assume that they satisfy Assumption 2.3 with $m > d/2$, $p = 2$ and $\vartheta = 0$. Thus by Sobolev's embedding of W^m_2 into C_b, the space of bounded continuous functions, for all ω we can find a continuous function of x which is equal to ψ almost everywhere, and for each t and ω we have continuous functions of x which coincide with f_t and g_t for almost every $x \in \mathbb{R}^d$. Here and in the following we always take such continuous modifications if they exist, thus we always assume that ψ, f_t, and g_t are continuous in x for all t (for $g = (g^k)^\infty_{k=1}$ this means, as usual, continuity as a function with values in l_2). In particular, terms like $f_t(x)$ in (3.1) make sense. We note that for $m > d/2$ one can use Sobolev's theorem on embedding W^m_2 to C_b to show also that if Assumption 2.3 holds with $m > d/2$, $p = 2$ $\theta = 0$, then

$$\sum_{x\in\mathbb{G}_h} |\psi(x)|^2 h^d + \int_0^T \Big(\sum_{x\in\mathbb{G}_h} |f_t(x)|^2 h^d + \sum_{x\in\mathbb{G}_h} \sum_k |g^k_t(x)|^2 h^d\Big)\, dt$$

$$\leq N\|\psi(x)\|_m^2 + N\int_0^T \|f_t\|_m^2 + \sum_k \|g_t^k\|_m^2\,dt < \infty\,\text{(a.s.)},$$

with a constant $N = N(\Lambda_0, d)$, where $\|\cdot\|_m := |\cdot|_{m,2,0}$. (See Lemma 4.2 in [10].)

Clearly, for $\varphi \in C_0^\infty(\mathbb{R}^d)$ and $\lambda \neq 0$

$$\delta_\lambda^h \varphi(x) \to \partial_\lambda \varphi(x) := \lambda^i D_i \varphi(x) \quad \text{as } h \to 0.$$

Thus, in order to approximate L and M^k by L^h and $M^{h,k}$, respectively, we need the following compatibility condition.

Assumption 3.1. *For every $i, j = 1, \ldots, d$, $k = 1, \ldots$ and for $P \otimes dt \otimes dx$-almost all $(\omega, t, x) \in \Omega \times [0,T] \times \mathbb{R}^d$*

$$a^{ij} = \sum_{\lambda,\kappa\in\Lambda_0} \mathfrak{a}^{\lambda\kappa}\lambda^i\kappa^j,\; b^i = \sum_{\lambda\in\Lambda_0}(\mathfrak{a}^{0\lambda} + \mathfrak{a}^{\lambda 0})\lambda^i,\; c = \mathfrak{a}^{00},$$

$$\sigma^{ik} = \sum_{\lambda\in\Lambda_0} \mathfrak{b}^{\lambda,k}\lambda^i, \quad \mu^k = \mathfrak{b}^{0,k}.$$

For each $x \in \mathbb{G}_h$ equation (3.1) is a stochastic differential equation (SDE), i.e., in general, (3.1)-(3.2) is an infinite system of SDEs. To replace this with a finite system we make the coefficients, together with the free and initial data, vanish outside of a large ball by multiplying them with a cutoff function ζ_R, which satisfies the following condition.

Assumption 3.2. *For an integer $m \geq 0$ and a real number $R > 0$ the function ζ_R is a continuous function with compact support on $\mathbb{R}^d$, such that $\zeta(x) = 1$ for $|x| \leq R$ and the derivatives of ζ_R up to order $m+1$ are continuous functions, bounded by a constant C.*

In this way we replace (3.1)-(3.2) with the system of SDEs

$$du_t(x) = (L_t^{h,R}u_t(x) + f_t^R(x))\,dt + (M_t^{h,R,k}u_t(x) + g_t^{R,k}(x))\,dw_t^k, \quad t \in [0,T], \tag{3.4}$$

with initial condition

$$u_0(x) = \psi^R(x), \tag{3.5}$$

for $x \in \mathbb{G}_h \cap \operatorname{supp}\zeta_\mathrm{R}$, where $\operatorname{supp}\zeta_\mathrm{R}$ is the support of ζ_R,

$$(\psi^R, f_t^R, g_t^{R,k}) := (\zeta_R\psi, \zeta_R f_t, \zeta_R g_t^k) \tag{3.6}$$

and

$$L_t^{h,R} := \sum_{\lambda,\kappa\in\Lambda_1} \mathfrak{a}_t^{\lambda\kappa,R}\delta_\lambda^h\delta_\kappa^h, \quad M_t^{h,R,k} := \sum_{\lambda\in\Lambda_1} \mathfrak{b}_t^{\lambda,R,k}\delta_\lambda^h, \quad k = 1, 2, \ldots,$$

with

$$\mathfrak{a}^{\lambda\kappa,R} := \zeta_R^2\mathfrak{a}^{\lambda\kappa} \quad \text{for } \lambda, \kappa \in \Lambda_0, \tag{3.7}$$

$$(\mathfrak{a}^{0\kappa,R}, \mathfrak{a}^{\lambda 0,R}, \mathfrak{b}^{\lambda,R,k}) := (\zeta_R\mathfrak{a}^{0\kappa}, \zeta_R\mathfrak{a}^{\lambda 0}, \zeta_R\mathfrak{b}^{\lambda,k}) \quad \text{for } \lambda, \kappa \in \Lambda_1,\ k \geq 1. \tag{3.8}$$

At this point our approximation is a finite dimensional (affine) linear SDE, whose coefficients are bounded by K owing to (3.3), and furthermore, by virtue of Remark 3.1, for each x, $f^R(x)$ and $g^R(x)$ are square integrable in time under Assumption 3.2 and 3.5 below with $m > d/2$.

Hence (3.4)-(3.5) has a unique solution

$$\{u_t^{h,R}(x) : x \in \mathbb{G}_h \cap \operatorname{supp} \zeta_{\mathrm{R}}\}_{\mathrm{t}\in[0,\mathrm{T}]},$$

by virtue of a well-known theorem of Itô on finite dimensional SDEs with Lipschitz continuous coefficients. The approximation of such equations are well studied, various time-discretization methods can be used, each of them with their own advantages and disadvantages. Here we chose the implicit Euler method, formulated as follows.

We take a mesh-size $\tau = T/n$ for an integer $n \geq 1$, and approximate (3.1)-(3.2) by the equations

$$v_0(x) = \psi^R(x), \tag{3.9}$$

$$v_i(x) = v_{i-1}(x) + (L^{h,R}_{\tau i} v_i(x) + f^R_{\tau(i-1)}(x))\tau$$
$$+ \sum_{k=1}^{\infty} (M^{h,R,k}_{\tau(i-1)} v_{i-1}(x) + g^{R,k}_{\tau(i-1)}(x))\xi^k_i, \quad i = 1, 2, \ldots, n, \tag{3.10}$$

for $x \in \mathbb{G}_h \cup \operatorname{supp} \eta_{\mathrm{R}}$, where $\xi^k_i = w^k_{i\tau} - w^k_{(i-1)\tau}$.

Remark 3.2. In many applications, including the Zakai equation for nonlinear filtering, the driving noise is finite dimensional. If this is not the case, one needs another level of approximation, at which the infinite sum in (3.10) is replaced by its first m terms. We shall not discuss this here.

Remark 3.3. As mentioned before, Euler approximations for SDEs are very well studied. Therefore, while it is far from immediate that the error is of the desired order, independently of h and R, the implementation of the scheme goes as usual, see e.g. [14] and its references.

To prove to solvability of the fully discretized equation (3.9)-(3.10) and estimate its error from the true solution of (2.1)-(2.2) on the space-time grid we pose the following assumptions. As for the following we confine ourselves to the L_2-scale, without weights, we use the shorthand notation $\|\cdot\|_m = |\cdot|_{W^m_{2,0}}$, $\|\cdot\| = \|\cdot\|_0$.

Assumption 3.3. *For all* $(\omega, t, x) \in \Omega \times [0, T] \times \mathbb{R}^d$

$$\sum_{\lambda,\kappa\in\Lambda_0} (2\mathfrak{a}^{\lambda\kappa} - \mathfrak{b}^{\lambda,k}\mathfrak{b}^{\kappa,k}) z^\lambda z^\kappa \geq 0$$

for all $z = (z^\lambda)_{\lambda\in\Lambda_0}$, $z^\lambda \in \mathbb{R}$.

In the following assumptions m and l are nonnegative integers, as before, and will be more specified in the theorems below.

Assumption 3.4. *The derivatives in* x *of* $\mathfrak{a}^{\lambda\kappa}$ *up to order* $\max(m, 2)$ *are* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}^d)$*-measurable functions, bounded by* K *for all* $\lambda, \kappa \in \Lambda_1$. *The derivatives in* x *of* $\mathfrak{b}^\lambda = (\mathfrak{b}^{\lambda r})_{r=1}^\infty$ *up to order* $m + 1$ *are* $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^d)$*-measurable* l_2*-valued functions, bounded by* K*, for all* $\lambda \in \Lambda_1$.

Assumption 3.5. *The initial value,* ψ *is an* $\mathcal{F}_0$*-measurable random variable with values in* W^m_2. *The free data,* f_t *and* $g_t = (g^k)_{k=1}^\infty$ *are predictable processes with values in* W^m_2 *and* $W^{m+1}_2(l_2)$*, respectively, such that almost surely*

$$\mathcal{K}^2_m := \|\psi\|^2_m + \int_0^T \left(\|f_t\|^2_m + \|g_t\|^2_{m+1}\right) dt < \infty. \tag{3.11}$$

Assumption 3.6. *There exists a constant H such that*

$$\mathbb{E}\|f_t - f_s\|_l^2 + \mathbb{E}\|g_s - g_t\|_{l+1}^2 \le H|t-s|, \quad \mathbb{E}\|f_t\|_{l+1}^2 + \mathbb{E}\|g_t\|_{l+2}^2 \le H$$

for all $s, t \in [0,T]$, and

$$|D^\alpha(\mathfrak{a}_t^{\lambda\kappa}(x) - \mathfrak{a}_s^{\lambda\kappa}(x))|^2 \le H|t-s|, \quad \sum_k |D^\beta(\mathfrak{b}_t^{\lambda,k}(x) - \mathfrak{b}_s^{\lambda,k}(x))|^2 \le H|t-s|,$$

for all $\omega \in \Omega$, $x \in \mathbb{R}^d$, $s,t \in [0,T]$ and multi-indices α and β with $|\alpha| \le l$ and $|\beta| \le l+1$.

Remark 3.4. If Assumptions 3.1 and 3.3 hold then

$$(2a^{ij} - \mu^{ir}\mu^{jr})z^i z^j = (2\sum_{\lambda,\kappa\in\Lambda_0} \mathfrak{a}^{\lambda\kappa}\lambda^i\kappa^j - \sum_{\lambda\in\Lambda_0}\mathfrak{b}^{\lambda,k}\lambda^i \sum_{\kappa\in\Lambda_0}\mathfrak{b}^{\kappa,k}\kappa^j)z^i z^j$$

$$= \sum_{\lambda,\kappa\in\Lambda_0}(2\mathfrak{a}^{\lambda\kappa} - \mathfrak{b}^{\lambda,k}\mathfrak{b}^{\kappa,k})(\lambda^i z^i)(\kappa^j z^j) \ge 0$$

for all $z = (z^1, ..., z^d) \in \mathbb{R}^d$, i.e., Assumption 2.1 also holds. Clearly, Assumptions 3.1 and 3.4 imply Assumption 2.2 (a)-(b). Notice that Assumption 3.5 is the same as Assumption 2.3 with $p = 2$ and $\vartheta = 0$. Thus if Assumptions 2.2 (c), and 3.1 through 3.5 hold with $m > 2 + d/2$, then by virtue of Corollary 2.5 equation (2.1) with initial condition (2.2) has a unique classical solution

$$u = \{u_t(x) : t \in [0,T], x \in \mathbb{R}^d\}.$$

Now we are in the position to formulate the first main theorem of this section, with the notation $\mathbb{G}_h^R = \mathbb{G}_h \cap B_R$.

Theorem 3.5. *Let $l > d/2$ be an integer. Let Assumptions 3.1 through 3.4 hold with $m \ge 4+l$, and let Assumptions 2.2 (c) and Assumption 3.6 hold with $m \ge 2+l$ and with $l+1$, respectively. Then if τ is sufficiently small, then for any $h > 0$, $R > 1$ the system of equations (3.9)-(3.10) has a unique solution $(v_i^{R,h,\tau})_{i=0}^n$. Moreover, for any $\nu \in (0,1)$, $q > 1$ we have*

$$\mathbb{E}\max_{i=0,\dots,n}\max_{x\in\mathbb{G}_h^{\nu R}}|u_{\tau i}(x) - v_i^{h,R,\tau}(x)|^2$$

$$\le N_1 e^{-\delta R^2}\mathbb{E}^{1/q}\mathcal{K}_{l+2}^{2q} + N_2(h^4+\tau)(1+\mathbb{E}\mathcal{K}_m^2), \tag{3.12}$$

with constants N_1 and $\delta > 0$ depending only on K, d, T, C q, ν and Λ_0, and a constant $N_2 = N_2(K,T,d,C,H,\Lambda_0)$.

As mentioned above we want to have approximations with higher order accuracy in h by extrapolating from $v^{h,R,\tau}$. Let us recall the method of Richardson's extrapolation. This technique, first introduced in [22], allows one to accelerate the rate of convergence by appropriately mixing approximations with different mesh sizes, given that a power expansion of the error in terms of the mesh sizes is available. We shall use this, based on results of [7], to obtain higher order approximations with respect to the spatial mesh size h. To formulate the extrapolation, let $r \ge 0$, V be the $(r+1)\times(r+1)$ Vandermonde matrix $V^{ij} = (4^{-(i-1)(j-1)})$,

$$(c_0, c_1, \dots, c_r) := (1, 0, \dots, 0)V^{-1}, \tag{3.13}$$

and define

$$\bar{v}^{h,R,\tau} := \sum_{i=0}^{r} c_i v^{h/2^i,R,\tau}, \tag{3.14}$$

where $v^{h/2^i,R,\tau}$ denotes the solution of (3.9)-(3.10) with $h/2^i$ in place of h. As we shall see, even by mixing only two approximations with different mesh sizes, that is, setting $r = 1$, the extrapolation increases the order of accuracy in h from 2 to 4.

The second main result of this section is the following.

Theorem 3.6. *In additions to the assumptions of Theorem 3.5 let Assumptions 3.2, 3.4 and 3.5 hold with $m \geq 4r + 4 + l$. Then for the extrapolation $\bar{v}^{h,R,\tau}$ we have*

$$\mathbb{E} \max_{i=0,\dots,n} \max_{x\in\mathbb{G}_h^{\nu R}} |u_{\tau i}(x) - \bar{v}_i^{h,R,\tau}(x)|^2$$

$$\leq N_1 e^{-\delta R^2} \mathbb{E}^{1/q} \mathcal{K}_{2+l}^{2q} + N_2(h^{2(2r+2)} + \tau)(1 + \mathbb{E}\mathcal{K}_m^2) \tag{3.15}$$

for any $\nu \in (0,1)$ and $q > 1$, with constants N_1 and $\delta > 0$, depending only on K, d, T, C, ν, q and Λ_0, and a constant $N_2 = N_2(K, T, d, C, H, r, \Lambda_0)$.

These theorems will be proved by using Theorem 2.6, some results from [7], summarized below in Theorem 3.9, and the error estimate for the time-discretization, established in Theorem 3.10 below.

Example 3.7. Consider the equation

$$du_t(x) = \sin^2(x) D^2 u_t(x)dt + \sin(x) Du_t(x)dw_t$$

for $(t,x) \in [0,1] \times \mathbb{R}$, where $(w_t)_{t\in[0,1]}$ is a 1-dimensional Wiener process, with the initial condition

$$u_0(x) = (1 + x^2)^{-1}.$$

The choice of localizing function ζ_R is quite arbitrary, for the sake of concreteness we take $\zeta_R(x) := f(x + 2 + R) - f(x - 2 - R)$, where

$$f(x) := \frac{2}{\pi} \arctan e^{x/(1-x^2)} \quad \text{for } |x| < 1, \text{ and } f(x) := \mathbf{1}_{[1,\infty)}(x) \text{ for } |x| \geq 1,$$

while noting that in practice a simple mollified indicator of $[-R, R]$ may be more favourable. Notice that $\zeta_R(x) = 1$ for $|x| \leq R$ and $\operatorname{supp} \zeta_R = [-3 - R, 3 + R]$.

For an integers $j \geq 1$ and $n \geq 1$ we set $h = R/(10j)$ and $\tau = 1/n$. To use the extrapolation with $r = 1$ in (3.14), we need to solve two discrete equations with spatial mesh sizes $\bar{h} = h, h/2$, and mix them according to (3.14), where one can check that the coefficients are $c_0 = -1/3, c_1 = 4/3$. Following the steps outlined above, the discrete equation we arrive at is

$$u_i^{R,\bar{h},\tau}(k\bar{h}) = \mathfrak{a}(k\bar{h})(\delta^{\bar{h}}\delta^{\bar{h}} u_i^{R,\bar{h}})(k\bar{h})\tau + \mathfrak{b}(k\bar{h})(\delta^{\bar{h}} u_{i-1}^{R,\bar{h}})(k\bar{h})(w_{\tau i} - w_{\tau(i-1)})$$

for $i = 1, 2, \dots, n$ and $k = 0, \pm 1, \dots, \pm\lceil (3 + R)/\bar{h}\rceil$, with the initial values

$$u_0^{R,\bar{h},\tau}(k\bar{h}) = (1 + k^2\bar{h}^2)^{-1}\zeta_R(k\bar{h}),$$

where $\mathfrak{a}(x) = \zeta_R^2(x) \sin^2(x)$, $\mathfrak{b}(x) = \zeta_R(x) \sin(x)$, and

$$\delta^{\bar{h}}\phi(x) = (2\bar{h})^{-1}[\phi(x + \bar{h}) - 2\phi(x) + \phi(x - \bar{h})].$$

For each $\bar{h}$ one can solve the above equation recursively in i.

Taking the resulting solutions $u^{R,h,\tau}$ and $u^{R,h/2,\tau}$, and setting

$$v^{R,h,\tau} = (4/3)u^{R,h/2,\tau} - (1/3)u^{R,h,\tau},$$

we can conclude that the error $\mathbb{E}\max_{i,k}|u_{i\tau}(kh) - v_i^{R,h,\tau}(kh)|^2$, where i runs over $0, 1, \dots, n$ and k runs over $0 \pm 1, \dots, \pm 0.9R/h$, is of order $e^{-\delta R^2} + h^8 + \tau$.

Before summarising some results from [7] on finite difference operators L^h, M^h and stochastic finite difference schemes in spatial variables we need to make an important remark.

Remark 3.8. The concept of a solution of (3.1)-(3.2), as a process with values in $l_{2,h}$, the space of functions $\phi : \mathbb{G}_h \to \mathbb{R}$ with finite norm $\|\phi\|^2_{l_{2,h}} = \sum_{x\in\mathbb{G}_h}|\phi(x)|^2$, is straightforward. One can, however, also consider (3.1)-(3.2) on the whole space, that is, for $(t,x) \in H_T$. Namely, when Assumptions 3.4 and 3.5 hold, then we can look for an $\mathcal{F}_t$-adapted L_2-valued solution $(u_t^h)_{t\in[0,T]}$ such that almost surely for every $t \in [0,T]$

$$u_t^h = \psi + \int_0^t (L_s^h u_s^h + f_s)\, ds + \sum_{k=1}^{\infty}\int_0^t (M_s^{h,k}u_s^h + g_s^k)\, dw_s^k \tag{3.16}$$

in the Hilbert space L_2, where the first integral is understood as Bochner integral of L_2-valued functions, the stochastic integrals are understood as Itô integrals of L_2-valued processes, and the convergence of their infinite sum is understood in probability, uniformly in $t \in [0,T]$. Thus, by a well-known theorem on SDEs in Hilbert spaces with Lipschitz continuous coefficients, equation (3.16) has a unique L_2-valued continuous $\mathcal{F}_t$-adapted solution $u^h = (u_t^h)_{t\in[0,T]}$. We refer to such a solution as an L_2*-valued solution* to (3.1)-(3.2). We can view equation (3.16) also as an SDE in the Hilbert space W_2^m, and by the same theorem on existence and uniqueness of solution to SDEs in Hilbert spaces, we get a unique W_2^m-valued continuous $\mathcal{F}_t$-adapted solution to it. Consequently, if Assumptions 3.4 and 3.5 hold with $m \geq 0$, then $u = (u_t^h)_{t\in[0,T]}$, the L_2-valued solution to (3.1)-(3.2) is a W_2^m-valued continuous $\mathcal{F}_t$-adapted process. Also, it is straightforward to see that these two concepts of solutions are "compatible" in the sense that if $m > d/2$, then the restriction of the L_2-valued solution to $\mathbb{G}_h$ solves (3.1)-(3.2) as an $l_{2,h}$-valued process. Note that (3.4)-(3.5) is a special case of the class of equations of the form (3.1)-(3.2), and so the above discussion applies to it as well.

The analogous concepts will be used for solutions of (3.9)-(3.10). Namely, a sequence of L_2-valued random variables $(v_i^{h,R})_{i=0}^n$ is called an L_2-valued solution to (3.9)-(3.10) if $v_i^{h,R}$ is $\mathcal{F}_{i\tau}$-measurable for $i = 0, 1, 2, ..., n$, and the equalities hold for almost all $x \in \mathbb{R}^d$, for almost all $\omega \in \Omega$. If $(v_i^{h,R})_{i=0}^n$ is an L_2-valued solution to (3.9)-(3.10) such that $v_i^{h,R} \in W_2^m$ (a.s.) for $m > d/2$ for each i, then it is easy to see that the restriction of the continuous version of $v_i^{h,R}$ to $\mathbb{G}_h \cap \operatorname{supp}\zeta_{\mathrm{R}}$ for each i gives a solution $\{v_i(x) : x \in \mathbb{G}_h \cap \operatorname{supp}\zeta_{\mathrm{R}}, \mathrm{i} = 0, 1, ..., \mathrm{n}\}$ to (3.9)-(3.10).

Theorem 3.9. *Let Assumptions 3.3, 3.4 and 3.5 hold with an integer $m \geq 0$. Then*
(a) For any $\phi \in W_p^m$ and $|\gamma| \leq m$

$$2(D^\gamma\phi, D^\gamma L_t^h\phi) + \sum_k \|D^\gamma M_t^{h,k}\phi\|^2 \leq N\|\phi\|_m^2$$

for all $\omega \in \Omega$ and $t \in [0,T]$ with a constant $N = N(K, m, d, \Lambda_0)$;

(b) *There is a unique L_2-valued solution u^h of* (3.1)-(3.2). *It is a W_2^m-valued process with probability one, and for any $q>0$*

$$\mathbb{E}\sup_{t\leq T}\|u_t^h\|_m^q \leq N\mathbb{E}\mathcal{K}_m^q$$

with a constant $N=N(K,m,d,\Lambda_0,q,T)$;

(c) *If for an integer $r\geq 0$ Assumptions 3.1 through 3.5 with $m>4r+4+d/2$ hold, then for any $q>0$*

$$\mathbb{E}\sup_{t\leq T}\max_{x\in\mathbb{G}_h}|u_t(x)-\bar{u}^h(x)|^q \leq Nh^{q(2r+2)}\mathbb{E}\mathcal{K}_m^q,$$

where u is the classical solution to (2.1)-(2.2), *$\bar{u}^h=\sum_{i=0}^r u^{h/2^i}$, and N is a constant depending only on K, d, T, q, Λ_0 and m.*

As discussed above, under Assumptions 3.4 and 3.5 we have a unique L_2-valued solution $u^{h,R}=(u_t^{h,R})_{t\in[0,T]}$ to (3.4)-(3.5), and it is a continuous W_2^m-valued process.

Theorem 3.10. *(i) Let Assumptions 3.2 through 3.5 hold with $m\geq 0$. Then for sufficiently small τ there exists for all h and $R>0$ a unique L_2-valued solution $v^{h,R,\tau}$ to* (3.9)-(3.10) *such that $v_i^{h,R,\tau}\in W_2^m$ for every $i=0,1,...,n$ and $\omega\in\Omega$.*
(ii) If Assumption 3.6 holds with some integer $l\geq 0$ and Assumptions 3.1 through 3.5 hold with $m=l+3$, then

$$\max_{i\leq n}\mathbb{E}\|u_{\tau i}^{h,R}-v_i^{h,R,\tau}\|_l^2 \leq N\tau(1+\mathbb{E}\mathcal{K}_m^2), \tag{3.17}$$

with a constant $N=N(K,C,H,d,T,l,\Lambda_0)$.
(iii) Let $l\geq 0$ be an integer. If Assumption 3.6 holds with $l+1$ in place of l, and Assumptions 3.1 through 3.5 hold with $m=l+4$, then

$$\mathbb{E}\max_{i\leq n}\|u_{\tau i}^{h,R}-v_i^{h,R,\tau}\|_l^2 \leq N\tau(1+\mathbb{E}\mathcal{K}_m^2) \tag{3.18}$$

with a constant $N=N(K,C,H,d,T,l,\Lambda_0)$.

Proof. To prove solvability of the system of equations (3.9)-(3.10) we rewrite (3.10) in the form

$$(I-\tau L_{\tau i}^{h,R})v_i = v_{i-1}+\tau f_{\tau(i-1)}+\sum_{k=1}^{\infty}(M_{\tau(i-1)}^{h,R,k}v_{i-1}+g_{\tau(i-1)}^{R,k})\xi_i^k, \quad i=1,2,\dots,n, \tag{3.19}$$

where I denotes the identity operator. We are going to show by induction on $j\leq n$ that for sufficiently small τ for each j there is a sequence of W_2^m-valued random variables $(v_i)_{i=0}^j$, such that v^i is $\mathcal{F}_{i\tau}$-measurable, $v_0=\psi^R$ and (3.19) holds for $1\leq i\leq j$. For $j=0$ there is nothing to prove. Let $j\geq 1$ and assume that our statement holds for $j-1$. Consider the equation

$$\mathbb{D}v=X, \tag{3.20}$$

where

$$\mathbb{D}:=I-\tau L_{\tau j}^{h,R}, \quad X:=v_{j-1}+\tau f_{\tau(j-1)}+\sum_{k=1}^{\infty}(M_{\tau(j-1)}^{h,R,k}v_{i-1}+g_{\tau(j-1)}^{R,k})\xi_j^k.$$

In the following we take κ to be either 0 or m. It easy to see that $\mathbb{D}$ is a bounded linear operator from W_2^κ into W_2^κ, for each $\omega\in\Omega$ and τ. Let us define the norm $|\cdot|_\kappa$ in W_2^κ by

$$|\varphi|_\kappa^2=\|(I-\Delta)^{\kappa/2}\phi\|^2=\sum_{\gamma:|\gamma|\leq\kappa}C_\gamma\|D^\gamma\phi\|^2,\quad \Delta:=\sum_{i=1}^d D_i^2,$$

where C_γ is a positive integer for each multi-index γ, $|\gamma|\leq\kappa$. Thus $|\cdot|_\kappa$ is a Hilbert norm which is equivalent to to $\|\cdot\|_\kappa$. We denote the corresponding inner product in W_2^κ by $(\ ,\ )_\kappa$. By virtue of Theorem 3.9 (a) for all $\omega\in\Omega$ and τ we have

$$(v,Dv)_\kappa=|v|_\kappa-\tau(L^{h,R}_{\tau j}v,v)_\kappa\geq|v|_\kappa^2-\tau N|v|^2_\kappa,\quad \text{for all } v\in W_2^\kappa$$

where the dependence of N is as in the theorem, in particular, it is independent of h,R. Consequently, for $\tau<1/N$ we have

$$(v,\mathbb{D}v)_\kappa\geq\delta|v|^2_\kappa\quad\text{for all } v\in W_2^\kappa,\quad \omega\in\Omega,$$

where $\delta=1-\tau N>0$. Hence by the Lax-Milgram lemma for every $\omega\in\Omega$ there is a unique $v=v_\kappa\in W_2^\kappa$ such that

$$(\mathbb{D}v_\kappa,\varphi)_\kappa=(X,\varphi)_\kappa\quad\text{for all } \varphi\in C_0^\infty(\mathbb{R}^d).$$

Since $(Y,\varphi)_\kappa=(Y,(I-\Delta)^\kappa\varphi)$ for all $Y\in W_2^\kappa$ and $\varphi\in C_0^\infty(\mathbb{R}^d)$, we have

$$(\mathbb{D}v_\kappa,(I-\Delta)^\kappa\varphi)=(X,(I-\Delta)^\kappa\varphi)\quad\text{for all } \varphi\in C_0^\infty(\mathbb{R}^d).$$

Hence, taking into account that $\{(I-\Delta)^m\varphi):\varphi\in C_0^\infty(\mathbb{R}^d)\}$ is dense in $W_2^0=L_2$, we get that v_m solves (3.20) in L_2 as well, so by uniqueness, $v_m=v_0$. This means (3.20) has a unique solution $v\in L_2$ for every $\omega\in\Omega$, and $v\in W_2^m$ for every $\omega\in\Omega$. Since X and $\mathbb{D}\phi$ are W_2^m-valued $\mathcal{F}_{j\tau}$-measurable random variables for every $\varphi\in W_2^m$, the unique solution $v\in W_2^m$ to (3.20) is also $\mathcal{F}_{j\tau}$-measurable. This finishes the induction, and the proof of statement (i) of the theorem.

For parts (ii) and (iii) $\mathbb{E}\mathcal{K}_m^2<\infty$ may and will be assumed. Hence by Theorem 3.9 (b),

$$\mathbb{E}\sup_{t\in[0,T]}|u_t^{h,R}|^2_\varrho\leq N\mathbb{E}\mathcal{K}^2_\varrho<\infty\quad\text{for } \varrho=0,1,2,...,m. \tag{3.21}$$

As Assumption 3.6 also holds, we have

$$\mathbb{E}\|\psi\|^2_{l+\kappa}+\max_{i\leq n}\mathbb{E}\|f_{t_i}\|^2_{l+\kappa}+\max_{i\leq n}\sum_k\mathbb{E}|g^k_{t_i}\|^2_{l+\kappa}<\infty \tag{3.22}$$

with $\kappa=0$ in part (ii) and $\kappa=1$ in part (iii), and hence, with the same κ,

$$\mathbb{E}\|v_i^{h,R,\tau}\|^2_{l+\kappa}<\infty\quad\text{for every } i=0,1,...,n. \tag{3.23}$$

To start the proof of (ii), let us fix a multi-index γ with length $|\gamma|=:\varrho\leq l$. From equations (3.9)-(3.10) and (3.1)-(3.2), we get that the error $e_i:=u^{h,R}_{\tau i}-v_i^{h,R,\tau}$ is

a W^m_2-valued $\mathcal{F}_{\tau i}$-measurable random variable, $i=0,\dots,n$, and $D^\gamma e_i$ is the L_2-valued solution of the equation

$$\begin{aligned}D^\gamma e_i =& D^\gamma e_{i-1}\\ &+ D^\gamma L^{h,R}_{\tau i} e_i \tau + \int_{\tau(i-1)}^{\tau i} D^\gamma F_s\,ds\\ &+ D^\gamma M^{h,R,k}_{\tau(i-1)} e_{i-1}\xi^k_i + \int_{\tau(i-1)}^{\tau i} D^\gamma G^k_s\,dw^k_s\end{aligned}$$

for $i=1,\dots,n$, with zero initial condition, where

$$F_t := L^{h,R}_t u^{h,R}_t - L^{h,R}_{\kappa_2(t)} u^{h,R}_{\kappa_2(t)} + f^R_t - f^R_{\kappa_1(t)}, \quad \kappa_2(t) := \kappa^n_2(t) := (\lfloor nt\rfloor + 1)/n,$$

$$G^k_t := M^{h,k}_t u^{h,R}_t - M^{h,R,k}_{\kappa_1(t)} u^{h,R}_{\kappa_1(t)} + g^{R,k}_t - g^k_{\kappa_1(t)}, \quad \kappa_1(t) := \kappa^n_1(t) := \lfloor nt\rfloor/n.$$

To ease notation we set

$$\mathfrak{L}_i := L^{h,R}_{\tau i}, \quad \mathfrak{M}^k_i := M^{h,R,k}_{\tau i}, \quad \mathfrak{F}_i := \int_{\tau(i-1)}^{\tau i} D^\gamma F_s\,ds, \quad \mathfrak{G}_i := \int_{\tau(i-1)}^{\tau i} D^\gamma G^k_s\,dw^k_s,$$

and by using the simple identity $\|b\|^2 - \|a\|^2 = 2(b,b-a) - \|b-a\|^2$ with $b := D^\gamma e_i$ and $a := D^\gamma e_{i-1}$, we get

$$\begin{aligned}&\|D^\gamma e_i\|^2 - \|D^\gamma e_{i-1}\|^2\\ &= 2(D^\gamma e_i, D^\gamma \mathfrak{L}_i e_i\tau + \mathfrak{F}_i) + 2(D^\gamma e_i, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i) - \|D^\gamma e_i - D^\gamma e_{i-1}\|^2\\ &= 2(D^\gamma e_i, D^\gamma \mathfrak{L}_i e_i\tau + \mathfrak{F}_i) + 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i)\\ &\quad + 2(D^\gamma e_i - D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i) - \|D^\gamma e_i - D^\gamma e_{i-1}\|^2\\ &= 2(D^\gamma e_i, D^\gamma \mathfrak{L}_i i e_i\tau + \mathfrak{F}_i) + 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i)\\ &+ \|D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i\|^2 - \|D^\gamma \mathfrak{L}_i e_i\tau + \mathfrak{F}_i\|^2\\ &\le I^{(1)}_i + I^{(2)}_i + I^{(3)}_i + I^{(4)}_i + I^{(5)}_i + I^{(6)}_i\end{aligned}\tag{3.24}$$

with

$$\begin{aligned}I^{(1)}_i :=& 2\tau(D^\gamma e_i, D^\gamma \mathfrak{L}_i e_i)\\ I^{(2)}_i :=& 2(D^\gamma e_i, \mathfrak{F}_i)\\ I^{(3)}_i :=& 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i + \mathfrak{G}_i)\\ I^{(4)}_i :=& \|D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i\|^2\\ I^{(5)}_i :=& 2(D^\gamma \mathfrak{M}^k_{i-1} e_{i-1}\xi^k_i, \mathfrak{G}_i)\\ I^{(6)}_i :=& \|\mathfrak{G}_i\|^2.\end{aligned}$$

By the Young and Jensen inequalities, and basic properties of stochastic Itô integrals we have

$$I^{(2)}_i \le \tau\|D^\gamma e_i\|^2 + \tau^{-1}\|\mathfrak{F}_i\|^2 \le \tau\|D^\gamma e_i\|^2 + \int_{\tau(i-1)}^{\tau i} \|F_s\|^2_\varrho\,ds, \quad \mathbb{E}I^{(3)}_i = 0,\tag{3.25}$$

$$\mathbb{E}I_i^{(4)} = \tau\mathbb{E}\sum_k \|D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}\|^2, \tag{3.26}$$

$$\mathbb{E}I_i^{(6)} = \mathbb{E}\int_{\tau(i-1)}^{\tau i} \sum_k \|D^\gamma G_s^k\|^2\,ds \le \mathbb{E}\int_{\tau(i-1)}^{\tau i} \sum_k \|G_s^k\|_\varrho^2\,ds. \tag{3.27}$$

By Itô's identity for stochastic integrals

$$\mathbb{E}I_i^{(5)} = 2\mathbb{E}\int_{\tau(i-1)}^{\tau i} \int_{\mathbb{R}^d} D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}(x) D^\gamma G_s^k(x)\,dx\,ds.$$

Here by integration by parts we drop one derivative from $D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}$ on the term $D^\gamma G_s^k$, and then by the Cauchy-Schwarz-Bunyakovsky and Young inequalities we get

$$\mathbb{E}I_i^{(5)} \le \tau N\mathbb{E}\|e_{i-1}\|_\varrho^2 + N\mathbb{E}\int_{\tau(i-1)}^{\tau i} \sum_k \|G_s^k\|_{\varrho+1}^2\,ds.$$

Using (3.26), by Theorem 3.9 (a) we have

$$\mathbb{E}I_i^{(1)} + \mathbb{E}I_i^{(4)} \le N\tau\mathbb{E}\|e_i\|_\varrho^2. \tag{3.28}$$

Therefore, by taking expectations and summing up (3.24) over i from 1 to $j \le n$, and over γ for $|\gamma| \le l$, we get

$$\mathbb{E}\|e_j\|_l^2 \le N_0\tau \sum_{i=1}^{j} \mathbb{E}\|e_i\|_l^2 + N_0\mathbb{E}\int_0^T (\|F_s\|_l^2 + \sum_k \|G_s^k\|_{l+1}^2)\,ds \tag{3.29}$$

for $j = 1, ..., n$, where $N_0 = N_0(K, C, \Lambda_0, l, d)$ is a constant. Notice that due to (3.23) and (3.21)

$$\mathbb{E}\|e_i\|_l^2 < \infty \quad i = 1, 2, ..., n,$$

and due to (3.21) and Assumptions 3.4 and 3.6 we have

$$\mathbb{E}\int_0^T \|F_s\|_l^2 + \sum_k \|G_s^k\|_{l+1}^2\,ds < \infty.$$

Hence the right-hand side of inequality (3.29) is finite. Thus when $\tau < 1/N_0$, from (3.29) by discrete Gronwall's lemma it follows that

$$\mathbb{E}\|e_j\|_l^2 \le N_0(1 - N_0\tau)^{-j}\mathbb{E}\int_0^T (\|F_t\|_l^2 + \sum_k \|G_t^k\|_{l+1}^2)\,dt \tag{3.30}$$

for $j = 1, ..., n$. Now we are going to show that

$$\mathbb{E}\|F_t\|_l^2 + \sum_k \mathbb{E}\|G_t^k\|_{l+1}^2 \le N\tau(\mathbb{E}\mathcal{K}_{l+3}^2 + 1) \tag{3.31}$$

for all $t \in [0, T]$ with a constant $N = N(K, C, c, d, T, l, \Lambda_0)$. To estimate $\mathbb{E}\|F_t\|_\varrho^2$, first notice that due to Assumption 3.6,

$$\mathbb{E}\|f_t^R - f_{\kappa_1(t)}^R\|_l^2 \le N\tau, \tag{3.32}$$

and due to Assumptions 3.2, 3.4 and 3.6

$$\|L_t^{h,R} u_t^{h,R} - L_{\kappa_2(t)}^{h,R} u_{\kappa_2(t)}^{h,R}\|_l^2 \le 2A_1(t) + 2A_2(t)$$

with

$$A_1(t) := \|(L_t^{h,R} - L_{\kappa_2(t)}^{h,R}) u_{\kappa_2(t)}^{h,R}\|_l^2 \le N\tau\|u_{\kappa_2(t)}^{h,R}\|_{l+2}^2,$$

$$A_2(t) := \|L_t^{h,R}(u_t^{h,R} - u_{\kappa_2(t)}^{h,R})\|_l^2 \le N\|u_{\kappa_2(t)}^{h,R} - u_t^{h,R}\|_{l+2}^2.$$

By virtue of Theorem (b) and Assumption 3.2,

$$\mathbb{E}A_1(t) \le N\tau \sup_{t\in[0,T]} \mathbb{E}\|u_t^{h,R}\|_{l+2}^2 \le N\tau\mathbb{E}\mathcal{K}_{l+2}^2. \tag{3.33}$$

To estimate A_2 we show that

$$\mathbb{E}\|u_t^{h,R} - u_s^{h,R}\|_{l+2}^2 \le N|t-s|(1+\mathbb{E}\mathcal{K}_{l+3}^2) \quad \text{for all } s,t\in[0,T]. \tag{3.34}$$

To this end we mollify $u^{h,R}$ in the spatial variable. We take a nonnegative $\phi \in C_0^\infty(\mathbb{R}^d)$ supported on the unit ball in $\mathbb{R}^d$ such that it has unit integral, and for functions φ, which are locally integrable in $x\in\mathbb{R}^d$, we define $\varphi^{(\varepsilon)}$ by

$$\varphi^{(\varepsilon)}(x) := \varepsilon^{-d}\int_{\mathbb{R}} \varphi(y)\phi((x-y)/\varepsilon)\,dy, \quad x\in\mathbb{R}^d, \quad \text{for } \varepsilon>0.$$

We will make use of the following known and easily verifiable properties of $\varphi^{(\varepsilon)}$. For integers $r\ge 0$ and $\varepsilon>0$,

$$\|\varphi^{(\varepsilon)} - \varphi\|_r \le \varepsilon\|\varphi\|_{r+1} \quad \text{for } \varphi\in W_2^{r+1}, \tag{3.35}$$

and

$$\|\varphi^{(\varepsilon)}\|_r \le \|\varphi\|_r, \quad \|(D_i\varphi)^{(\varepsilon)}\|_r = \|D_i\varphi^{(\varepsilon)}\|_r \le \frac{N_i}{\varepsilon}\|\varphi\|_r, \quad \varphi\in W_2^r \tag{3.36}$$

for $i=1,2,...,d$, where N_i depends only on the sup and L_1 norms of $D_i\phi$. Thus by (3.35)

$$\begin{aligned}
\|u_t^{h,R} - u_s^{h,R}\|_{l+2}^2 &= \|u_t^{h,R} - u_s^{h,R}\|_{l+1}^2 + \sum_{i=1}^d \|D_i(u_t^{h,R} - u_s^{h,R})\|_{l+1}^2 \\
&\le \|u_t^{h,R} - u_s^{h,R}\|_{l+1}^2 + \sum_{i=1}^d \|D_i(u_t^{h,R} - u_s^{h,R})^{(\varepsilon)}\|_{l+1}^2 \\
&+ \sum_{i=1}^d \|D_iu_t^{h,R} - (D_iu_t^{h,R})^{(\varepsilon)}\|_{l+1}^2 + \sum_{i=1}^d \|(D_iu_s^{h,R} - (D_iu_s^{h,R})^{(\varepsilon)}\|_{l+1}^2 \\
&\le \|u_t^{h,R} - u_s^{h,R}\|_{l+1}^2 + \sum_{i=1}^d \|D_i(u_t^{h,R} - u_s^{h,R})^{(\varepsilon)}\|_{l+1}^2 \\
&\quad + N\varepsilon^2(\|u_t^{h,R}\|_{l+2} + \|u_s^{h,R}\|_{l+2}).
\end{aligned} \tag{3.37}$$

Since $u^{h,R}$ satisfies (3.1)-(3.2), for $0\le s\le t\le T$ we have

$$\|u_t^{h,R} - u_s^{h,R}\|_{l+1}^2 \le 2B_1 + 2B_2 \tag{3.38}$$

with

$$B_1 := \Big\| \int_s^t (L_r^{h,R}u_r^{h,R} + f_r^R)\,dr\Big\|_{l+1}^2,$$

$$B_2 := \Big\| \int_s^t (M_r^{h,R,k}u_r^{h,R} + g_r^{R,k})\,dw_r^k\Big\|_{l+1}^2,$$

and using (3.36) we have

$$\sum_{i=1}^d \|D_i(u_t^{h,R} - u_s^{h,R})^{(\varepsilon)}\|_{l+1}^2 \le \frac{N}{\varepsilon^2}B_1 + B_3, \tag{3.39}$$

with

$$B_3 := \sum_{i=1}^{d} \Big\| \int_s^t (D_i M_r^{h,R,k} u_r^{h,R} + D_i g_r^{R,k})\, dw_r^k \Big\|_{l+1}^2 .$$

It is easy to see that

$$\begin{aligned}\mathbb{E}B_1 &\le (t-s)\int_s^t \mathbb{E}\|L_s^{h,R}u_s^{h,R} + f_s^R\|_{l+1}^2\, ds \\ &\le N(t-s)^2(\sup_{t\le T}\mathbb{E}\|u_t\|_{l+3}^2 + \sup_{t\le T}\mathbb{E}\|f_t\|_{l+1}^2). \end{aligned} \tag{3.40}$$

Using Itô's identity, for B_2 we obtain

$$\begin{aligned}\mathbb{E}B_2 =& \mathbb{E}\int_s^t \sum_k \|M_r^{h,R,k}u_r^{h,R} + g_r^k\|_{l+1}^2\, dr \\ &\le N(t-s)(\sup_{t\le T}\mathbb{E}\|u_t^{h,R}\|_{l+2}^2 + \sup_{t\le T}\mathbb{E}\|g_t\|_{l+1}^2), \end{aligned} \tag{3.41}$$

and in the same way, for B_3 we have

$$\mathbb{E}B_3 \le N(t-s)(\sup_{t\le T}\mathbb{E}\|u_t^{h,R}\|_{l+3}^2 + \sup_{t\le T}\mathbb{E}\|g_t\|_{l+2}^2). \tag{3.42}$$

From (3.37) through (3.41) we get

$$\mathbb{E}\|u_t^{h,R} - u_s^{h,R}\|_{l+2}^2 \le N(\varepsilon^2 + (t-s)^2\varepsilon^{-2} + |t-s|)J$$

for every $\varepsilon > 0$, where

$$J := \sup_{t\le T}\{\mathbb{E}\|u_t^{h,R}\|_{l+3}^2 + \mathbb{E}\|g_t\|_{l+2}^2 + \mathbb{E}\|f_t\|_{l+1}^2\}.$$

Choosing here $\varepsilon := |t-s|^{1/2}$ gives

$$\mathbb{E}\|u_t^{h,R} - u_s^{h,R}\|_{l+2}^2 \le 3N(t-s)J.$$

Hence we obtain (3.34), since by Theorem 3.9 and Assumption 3.6 we have

$$J \le N\mathbb{E}\mathcal{K}_{l+3}^2 + H.$$

From (3.34) we get the estimate

$$\mathbb{E}A_1 \le N\tau(\mathbb{E}\mathcal{K}_{l+3}^2 + 1),$$

which together with (3.33) and (3.32) shows that $\mathbb{E}\|F_t\|_l^2$ is estimated by the right-hand side of the inequality (3.31). Similarly, by making use of (3.34), we can get the same estimate for $\mathbb{E}\|G_t\|_{l+1}^2$, which finishes the proof of (3.31). From (3.30) and (3.31) we have

$$\max_{j\le n}\mathbb{E}\|e_j\|_\varrho^2 \le \tau N(1-N_0\tau)^{-n}(\mathbb{E}\mathcal{K}_{\varrho+3} + 1)$$

for $\tau < N_0^{-1}$, with a constant $N = N(K, C, H, d, l, T, \Lambda_0)$. Hence noticing that

$$\lim_{n\to\infty}(1-N_0\tau)^{-n} = e^{N_0T},$$

we obtain (3.17), which completes the proof of (ii).

To prove (iii) notice first that by using (ii) with $l+1$ in place of l we have

$$\mathbb{E}\sum_{i=1}^{n}\|e_i\|_{l+1}^2 \le N(1 + \mathbb{E}\mathcal{K}_m^2) \tag{3.43}$$

with $m = l+4$. Fixing a multi-index γ as in (ii), with $|\gamma| = \varrho \le l$, we revisit (3.24) and observe that $I_i^{(5)} \le I_i^{(4)} + I_i^{(6)}$ and by Theorem 3.9

$$I_i^{(1)} \le N\tau \|e_i\|_l^2.$$

Thus summing up (3.24) over $i = 1, ..., j$, and recalling also (3.25) and (3.26), we obtain

$$\mathbb{E} \max_{1\le j\le n} \|D^\gamma e_i\|^2 \le N\tau \mathbb{E} \sum_{i=1}^n \|e_i\|_l^2 + N \int_0^T \mathbb{E}\|F_s\|_l^2 \, ds$$

$$+N \int_0^T \mathbb{E}\|G_s\|_l^2 \, ds + \tau \sum_{i=1}^n \mathbb{E} \sum_k \|D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}\|^2 + \mathbb{E} \max_{1\le j\le n} \sum_{i=1}^j I_i^{(3)}.$$

Hence, noticing that

$$\sum_{k=1}^\infty \|D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}\|^2 \le N\|e_i\|_{l+1}^2,$$

by using (3.43) and (3.31), we get

$$\mathbb{E} \max_{1\le i\le n} \|D^\gamma e_i\|^2 \le N\tau(\mathbb{E}\mathcal{K}_m^2 + 1) + \mathbb{E} \max_{1\le j\le n} \sum_{i=1}^j I_i^{(3)}. \tag{3.44}$$

Clearly,

$$\sum_{i=1}^j I_i^{(3)} = \int_0^{j\tau} Z_t^k \, dw_t^k, \quad j = 1, 2, ..., n,$$

where $(Z_t^k)_{t\in[0,T]}$ is defined by

$$Z_t^k = 2(D^\gamma e_{i-1}, D^\gamma \mathfrak{M}_{i-1}^k e_{i-1} + D^\gamma G_t^k) \quad \text{for } t \in (t_{i-1}, t_i],\ i = 1, 2, ..., n.$$

It is easy to see that

$$\int_0^T \sum_k |Z_t^k|^2 \, dt \le 4 \max_{1\le i\le n} \|D^\gamma e_i\|^2 \sum_k \left(\sum_{i=1}^n \|D^\gamma \mathfrak{M}_{i-1}^k e_{i-1}\|^2 \tau + \int_0^T \|G_s^k\|_l^2 \, ds \right)$$

$$\le \frac{1}{36} \max_{1\le i\le n} \|D^\gamma e_i\|^4 + 144 \left(\sum_{i=1}^n \|e_{i-1}\|_{l+1}^2 \tau + \sum_k \int_0^T \|G_s^k\|_l^2 \, ds \right)^2.$$

Hence by the Davis inequality we have

$$S_n \le \mathbb{E} \sup_{t\in[0,T]} \int_0^t Z_s^k \, dw_s^k \le 3\mathbb{E} \left(\int_0^T \sum_k |Z_s^k|^2 \, ds \right)^{1/2}$$

$$\le \frac{1}{2} \mathbb{E} \max_{1\le i\le n} \|D^\gamma e_i\|^2 + 36\tau \sum_{i=1}^n \mathbb{E}\|e_{i-1}\|_{l+1}^2 + 36\mathbb{E} \int_0^T \|G_s\|_l^2 \, ds < \infty.$$

Using this, and estimates (3.31) and (3.43), we obtain (iii) from (3.44). □

By virtue of Remark 3.8, by using Sobolev's theorem on embedding W_2^m into C_b for $m > d/2$, we get the following corollary.

Corollary 3.11. (i) Let Assumptions 3.2 through 3.5 hold with $m > d/2$. Then for sufficiently small τ there exists for all h and $R > 0$ a unique solution $v^{h,R,\tau} = \{v^{h,R,\tau}_{i\tau}(x) : x \in \mathbb{G}_h \cap \operatorname{supp} \zeta_{\mathrm{R}}\}$ to (3.9)-(3.10).
(ii) If Assumption 3.6 holds with some integer $l > d/2$ and Assumptions 3.1 through 3.5 hold with $m > d/2+3$, then

$$\max_{0\le i\le n} \mathbb{E} \max_{x\in\mathbb{G}_h} |u^{h,R}_{\tau i} - v^{h,R,\tau}_i|^2 \le N\tau(1+\mathbb{E}\mathcal{K}^2_m), \tag{3.45}$$

with a constant $N = N(K, C, H, d, T, l, \Lambda_0)$.
(iii) If Assumption 3.6 holds with $l > d/2 + 1$ and Assumptions 3.1 through 3.5 hold with $m > d/2 + 4$, then

$$\mathbb{E} \max_{i\le n} \max_{x\in\mathbb{G}_h} |u^{h,R}_{\tau i} - v^{h,R,\tau}_i|^2 \le N\tau(1+\mathbb{E}\mathcal{K}^2_m) \tag{3.46}$$

with a constant $N = N(K, C, H, d, T, l, \Lambda_0)$.

Proof of Theorems 3.5 and 3.6. The solvability of (3.9)-(3.10) has already been discussed above, so we need only prove the estimates (3.12) and (3.15). A natural way to separate the errors of the different types of approximations would be to write

$$|u_{\tau i}(x) - v^{h,R,\tau}_i(x)| \le |u_{\tau i}(x) - u^h_{\tau i}(x)| + |u^h_{\tau i}(x) - u^{h,R}_{\tau i}(x)| + |u^{h,R}_{\tau i}(x) - v^{h,R,\tau}_i(x)|.$$

However, in such a decomposition we cannot directly estimate the middle term on the right-hand side, that is, the localization error of the finite difference equation. Therefore we introduce $u^{0,R}$, the classical solution of (2.1)-(2.2) with data

$$\bar{\mathfrak{D}} := \mathfrak{D}^R = (\zeta_R\psi, \zeta_R^2 a, \zeta_R b, \zeta_R c, \zeta_R\sigma, \zeta_R\mu, \zeta_R f, \zeta_R g).$$

Clearly the pair $\mathfrak{D}$, $\bar{\mathfrak{D}}$ satisfies Assumption 2.4. Also, as the finite difference coefficients $\mathfrak{a}, \mathfrak{b}$ are compatible with the data $\mathfrak{D}$ in the sense that Assumption 3.1 is satisfied, it follows that the coefficients $\mathfrak{a}^R, \mathfrak{b}^R$, as defined in (3.7)-(3.8), are compatible with the data $\bar{\mathfrak{D}}$ in the same sense. Therefore in the decomposition

$$|u_{\tau i}(x) - v^{h,R,\tau}_i(x)| \le |u_{\tau i}(x) - u^{0,R}_{\tau i}(x)| + |u^{0,R}_{\tau i}(x) - u^{h,R}_{\tau i}(x)| + |u^{h,R}_{\tau i}(x) - v^{h,R,\tau}_i(x)|$$

the first term can be treated by Theorem 2.6, the second by Theorem 3.9 (c), and the third by Corollary 3.11 (iii). Adding up the resulting errors we get the estimate (3.12), and the dependence of the constants also follows from the invoked theorems.

Similarly, for (3.15) we write

$$\begin{aligned} |u_{\tau i}(x) - \bar{v}^{h,R,\tau}_i(x)| \le{}& |u_{\tau i}(x) - u^{0,R}_{\tau i}(x)| + |u^{0,R}_{\tau i}(x) - \bar{u}^{h,R}_{\tau i}(x)| \\ &+ \sum_{j=0}^{r} c_j |u^{h/2^j,R}_{\tau i}(x) - v^{h/2^j,R,\tau}_i(x)|, \end{aligned}$$

where $\bar{u}^{h,R} = \sum_{j=0}^r c_j u^{h/2^j,R}$, and follow the same steps as above.

Remark 3.12. As it can be easily seen from the last step of the proof, Assumption 3.6 can be weakened to α-Hölder continuity for any fixed $\alpha > 0$, at the cost of lowering the rate from $1/2$ to $\alpha \wedge (1/2)$.

To decrease the spatial regularity conditions, in particular, the term $d/2$, one would need the generalization of the results of [7], and subsequently, of Theorem 3.10, to arbitrary Sobolev spaces W^m_p. Partial results in this direction can be found in [5].

Acknowledgements. The authors thank the referees for their remarks and suggestions which helped to improve the presentation of the paper.

References

1. A. Debussche, J. Printems. Weak order for the discretization of the stochastic heat equation, Math. Comp., 78 (2009), 845-863
2. H. Dong, N.V. Krylov, Rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients, Electron. J. Differential Equations (2005), No. 102, 25 pp
3. M.I. Freidlin, On the factorization of non-negative definite matrices, Theory Probab. Appl. 13 (1968), 354-356.
4. L. Gerencsér, On a class of mixing processes, Stochastics and Stochastic Reports, Vol 26 (1989), 165-191.
5. M. Gerencsér, I. Gyöngy, Finite difference schemes for stochastic partial differential equations in Sobolev spaces, Appl. Math. Optim. 72 (2015), no. 1, pp 77-100
6. M. Gerencsér, I. Gyöngy and N.V. Krylov, On the solvability of degenerate stochastic partial differential equations in Sobolev spaces, Stoch. Partial Differ. Equ. Anal. Comput. 3 (2015), no. 1, pp 52-83
7. I. Gyöngy, On stochastic finite difference schemes, Stoch. Partial Differ. Equ. Anal. Comput. 2 (2014), no. 4, pp 539-583
8. I. Gyöngy and N.V. Krylov, On the rate of convergence of splitting-up approximations for SPDEs, Progress in Probability, Vol. 56, 301-321, 2003 Birkhäuser Verlag.
9. I. Gyöngy and A. Millet, On discretization schemes for stochastic evolution equations, Pot. Anal. 23 (2005), no. 2, 99-134
10. I. Gyöngy and N.V. Krylov, Accelerated Finite Difference Schemes for Linear Stochastic Partial Differential Equations in the Whole Space, SIAM J. Math. Anal., 42 (2010), no. 5, 2275-2296.
11. E. Hausenblaus, Approximation for semilinear evolution equations, Pot. Anal. 18 (2003), no. 2, 141-186
12. N. Hilber, O. Reichmann, C. Schwab, C. Winter, *Computational Methods for Quantitative Finance: Finite Element Methods for Derivative Pricing*, Springer-Verlag, Berlin, 2013
13. A. Jentzen, P.E. Kloeden, Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive spacetime noise, Proc. R. Soc. A (2009) 465, 649-667
14. P.E. Kloeden, E. Platen, H. Schurz, *Numerical Solution of SDE Through Computer Experiments*, Springer-Verlag, Berlin-Heidelberg, 1994.
15. N.V. Krylov, *Controlled Diffusion Processes*, Springer-Verlag, Berlin, 1980.
16. N.V. Krylov, On the Itô-Wentzell formula for distribution-valued processes and related topics, Probab. Theory Relat. Fields 150 (2011), 295–319.
17. N.V. Krylov, A W_2^n-theory of the Dirichlet problem for SPDEs in general smooth domains, Probab. Theory Relat. Fields 98, 389-421 (1994)
18. N.V. Krylov and B.L. Rozovskii, Stochastic partial differential equations and diffusion processes, Russian Math. Surveys Vol. 37, 6 (1982), 81–105.
19. H. Kunita, *Stochastic Flows and Stochastic Differential Equations*, Cambridge University Press, 1997.
20. D. Lamberton and B. Lapeyre, *Introduction to stochastic calculus applied to finance*, Chapman & Hall/CRC Financial Mathematics Series, Taylor & Francis, 1996.
21. J-M. Leahy, R. Mikulevičius, On Classical Solutions of Linear Stochastic Integro-Differential Equations, http://arxiv.org/abs/1404.0345.
22. L.F. Richardson, The approximate arithmetical solution by finite differences of physical problems involving differential equations, Philos.Trans. Roy. Soc. London, Ser.A, 210 (1910), 307-357.
23. D. Siška, Error Estimates for Approximations of American Put Option Price, Computational Methods in Applied Mathematics Vol. 12 (2012), No. 1, 108–120.
24. J.B. Walsh, Finite Element Methods for Parabolic Stochastic PDEs, Potential Anal. 23 (2005), no. 1, 1-43

25. Y. Yan, Galerkin finite element methods for stochastic parabolic partial differential equations, SIAM J. Numer. Anal. 43 (2005), n. 4, 1363-1384
26. H. Yoo, Semi-discretization of stochastic partial differential equations on $\mathbb{R}^1$ by a finite-difference method. Math. Comp. 69 (2000), no. 230, 653-666

School of Mathematics and Maxwell Institute, The University of Edinburgh, The King's Buildings, Peter Guthrie Tait Road, EDINBURGH, EH9 3FD Scotland, UK
E-mail address: `m.gerencser@sms.ed.ac.uk`

School of Mathematics and Maxwell Institute, The University of Edinburgh, The King's Buildings, Peter Guthrie Tait Road, EDINBURGH, EH9 3FD Scotland, UK
E-mail address: `gyongy@maths.ed.ac.uk`